Limites fluides de quelques échantillonneurs MCMC

Gersende FORT

CNRS LTCI, ENST, Paris

Travail en collaboration avec

- Sean MEYN (Université de l'Illinois, Urbana, US).
- Eric MOULINES (Ecole Nationale Supérieure des Télécommunications, Paris).
- Pierre PRIOURET (LPMA, Université Paris VI, Paris).

Soit $\{\Phi_k\}_{k\in\mathbb{N}}$ une chaîne de Markov à valeur dans $(\mathsf{X}, \mathcal{X})$ $\mathsf{X} = \mathbb{R}^d$.

Etude de la stabilité des chaînes de Markov :

$$\sup_{x\in C} \mathbb{E}_x \left[\sum_{k=0}^{\tau_C - 1} \rho(k)\, f(\Phi_k) \right],$$

où

C est un *petite* set, $\tau_C = \inf\{k \geq 1, \Phi_k \in C\}$.

$f \geq 1$: fonction.

$\rho \geq 0$: fonction de taux.

Ex. $\rho(n) = \kappa^n \quad \kappa > 1, \quad \rho(n) = n^\alpha (\log n)^\beta \, \exp(\gamma n^\delta) \quad 0 < \delta < 1$.

Pourquoi ce critère ?

Cas : X discret $\qquad \pi(A) \propto \mathbb{E}_{x_*}\left[\sum_{k=0}^{\tau_{x_*}-1} \mathbb{1}_A(\Phi_k)\right]$

$$\lim_n \|P^n(x,\cdot)-\pi(\cdot)\|_{\mathrm{TV}} = 0 \quad \longleftrightarrow \quad \mathbb{E}_{x_*}\left[\tau_{x_*}\right] = \mathbb{E}_{x_*}\left[\sum_{k=0}^{\tau_{x_*}-1} 1 \times 1\right] < \infty.$$

Cas : X général $\qquad$ pour un ensemble C petite

$$\lim_n \|P^n(x,\cdot)-\pi(\cdot)\|_{\mathrm{TV}} = 0 \quad \longleftrightarrow \quad \sup_{x\in C} \mathbb{E}_x\left[\tau_C\right] = \sup_{x\in C} \mathbb{E}_x\left[\sum_{k=0}^{\tau_C-1} 1 \times 1\right] < \infty.$$

$$\lim_n \rho(n)\|P^n(x,\cdot)-\pi(\cdot)\|_{\mathrm{TV}} = 0 \quad \longleftrightarrow \quad \sup_{x\in C} \mathbb{E}_x\left[\sum_{k=0}^{\tau_C-1} \rho(k)\right] < \infty.$$

$$\lim_n \rho(n) \sup_{g,|g|\leq f} |\mathbb{E}_x[g(\Phi_n)] - \pi(g)| = 0 \quad \longleftrightarrow \quad \sup_{x\in C} \mathbb{E}_x\left[\sum_{k=0}^{\tau_C-1} \rho(k) f(\Phi_k)\right] < \infty.$$

Soit $\{\Phi_k\}_{k\in\mathbb{N}}$ une chaîne de Markov à valeur dans $(\mathsf{X}, \mathcal{X})$ $\quad \mathsf{X} = \mathbb{R}^d$.

Etude de la stabilité des chaînes de Markov :

$$\sup_{x\in C} \mathbb{E}_x \left[\sum_{k=0}^{\tau_C - 1} \rho(k)\ f(\Phi_k) \right],$$

où

C est un *petite* set, $\tau_C = \inf\{k \geq 1, \Phi_k \in C\}$.

$f \geq 1$: fonction.

$\rho \geq 0$: fonction de taux.

Ex. $\rho(n) = \kappa^n \quad \kappa > 1, \quad \rho(n) = n^\alpha (\log n)^\beta \ \exp(\gamma n^\delta) \quad 0 < \delta < 1$.

Applications : Etude de quelques échantillonneurs MCMC.

$\hookrightarrow$ Par l'étude de la limite fluide associée.

Etude du processus $r > 0, \beta \geq 0$

$$\eta_r^\beta(t,x) = \frac{1}{r}\Phi_{\lfloor tr^{1+\beta} \rfloor}, \qquad \eta_r^\beta(0,x) = \frac{1}{r}\Phi_0 = x.$$

i.e.

$$\eta_r^\beta(t,x) = \frac{1}{r}\Phi_k \quad \text{pout tout } \frac{k}{r^{1+\beta}} \leq t < \frac{k+1}{r^{1+\beta}}.$$

Exemple : chaîne de Hastings-Métropolis à marche aléatoire symétrique

▷ De mesure invariante sur $\mathbb{R}^2$ donnée par

$$\pi(x) \propto \left(1 + x_1^2 + x_2^2 + x_1^8 x_2^2\right) \; \exp(-(x_1^2 + x_2^2)).$$

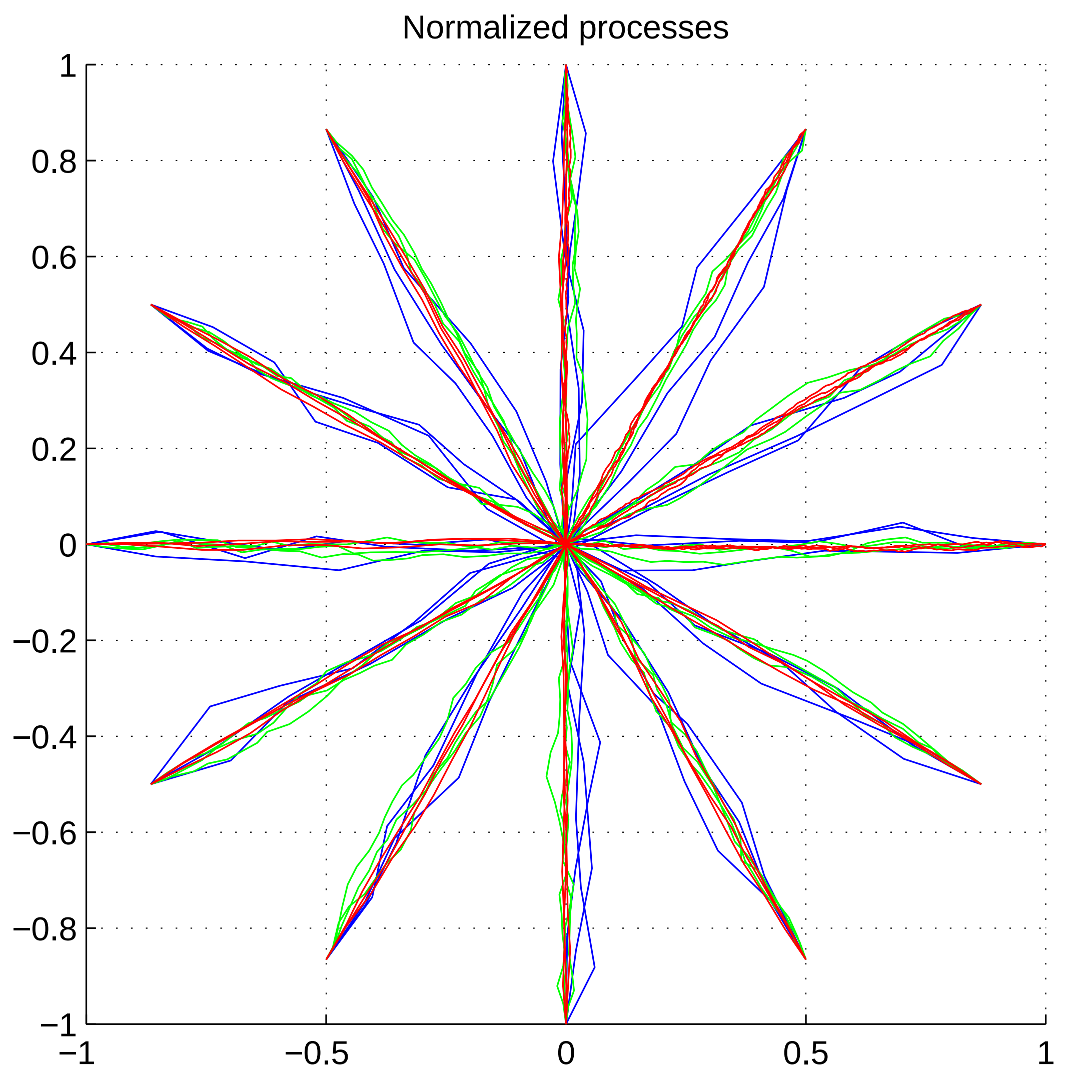
▷ Décrite par

$$\Phi_{k+1} = \Phi_k + Y_{k+1} \qquad \text{avec proba} \qquad \alpha(\Phi_k, \Phi_k + Y_{k+1}),$$

$$\Phi_{k+1} = \Phi_k \qquad \text{avec proba} \qquad 1 - \alpha(\Phi_k, \Phi_k + Y_{k+1}),$$

où

$$Y_{k+1} | \mathcal{F}_k \sim q(\Phi_k, \cdot), \qquad \alpha(x,z) = 1 \wedge \frac{\pi(z)}{\pi(x)}.$$



$\eta_{r,0}(\cdot, x)$, pour différentes valeurs initiales x sur la sphère unité et $r = 500$, 2000 et 10000. 3 trajectoires par couple (r, x).

$$\Phi_{k+1} = \Phi_k + \underbrace{\left(\mathbb{E}_x\left[\Phi_{k+1}|\mathcal{F}_k\right] - \Phi_k\right)}_{\Delta(\Phi_k)} + \underbrace{\left(\Phi_{k+1} - \mathbb{E}_x\left[\Phi_{k+1}|\mathcal{F}_k\right]\right)}_{\epsilon_{k+1}}.$$

$\eta_r^\beta(\cdot, x)$: processus cadlag $\longrightarrow$ loi $\mathbb{Q}^\beta_{r;x}$ sur l'espace des fonctions cadlag $\mathbb{R}^+ \to$ X.

Etant donné : $\{r_n\}_n \to +\infty \qquad \{x_n\}_n \to x$: $\qquad \mathbb{Q}^\beta_{r_n;x_n} \Rightarrow_n ?$

Supposons

- $\exists p > 1, \qquad \lim_K \sup_{x\in \mathsf{X}} \mathbb{E}_x\left[|\epsilon_1|^p 1_{|\epsilon_1|>K}\right] = 0.$
- $\exists\, \beta \in [0, 1 \wedge (p-1)), \qquad 0 < \sup_{x\in\mathsf{X}} (1 + |x|^\beta)|\Delta(x)| < \infty.$

Alors $\quad \forall\ 0 \leq \alpha \leq \beta$

- il existe des sous-suites $\{r_{n_j}\}_j, \{x_{n_j}\}_j$: $\mathbb{Q}^\alpha_{r_{n_j};x_{n_j}} \Rightarrow \mathbb{Q}^\alpha_x$.
- $\mathbb{Q}^\alpha_x$ est une probabilité sur C$(\mathbb{R}^+, \mathsf{X})$ fonctions continues $\mathbb{R}^+ \to$ X.
- pour $\alpha < \beta$, $\mathbb{Q}^\alpha_x = \delta_x$.

$\hookrightarrow \mathbb{Q}^\alpha_x$: α-limite fluide.

Limite fluide : Caractérisation

$$\Phi_{k+1} = \Phi_k + \underbrace{\left(\mathbb{E}_x\left[\Phi_{k+1} | \mathcal{F}_k\right] - \Phi_k\right)}_{\Delta(\Phi_k)} + \underbrace{\left(\Phi_{k+1} - \mathbb{E}_x\left[\Phi_{k+1} | \mathcal{F}_k\right]\right)}_{\epsilon_{k+1}}$$

Pour le processus normalisé : cas $\beta = 0$, pour simplifier

$$\begin{aligned} \eta_r\left[\frac{k+1}{r}, x\right] &= \frac{1}{r}\Phi_{k+1} \\ &= \eta_r\left[\frac{k}{r}, x\right] + \frac{1}{r}\Delta\left(r\,\eta_r\left[\frac{k}{r}, x\right]\right) + \frac{1}{r}\epsilon_{k+1} \\ &= \eta_r\left[\frac{k}{r}, x\right] + \frac{1}{r}\; h\left(\eta_r\left[\frac{k}{r}, x\right]\right) + \frac{1}{r}\left(\xi_k + \epsilon_{k+1}\right) \end{aligned}$$

en ayant posé

$$h\,(x) = \lim_{r\to\infty} \Delta(r\,x).$$

Ainsi version bruitée de

$$\mu\left(\frac{k+1}{r}\right) = \mu\left(\frac{k}{r}\right) + \frac{1}{r}h\left(\frac{k}{r}\right) \longleftrightarrow \text{ODE : } \dot{\mu}(t) = h(\mu(t))$$

Limite fluide : Caractérisation

Théorème :

Supposons qu'il existe

- un cône ouvert O de $\mathsf{X} \setminus \{0\}$
- une fonction continue $\Delta_\infty : \mathsf{O} \to \mathsf{X}$ t.q. pour tout compact $\mathsf{H} \subset \mathsf{O}$,

$$\lim_{r \to \infty} \sup_{x \in \mathsf{H}} \left| r^\beta |x|^\beta \Delta(rx) - \Delta_\infty(x) \right| = 0,$$

Alors pour tout $0 \leq s \leq t$, sur $\{\eta \in \mathsf{C}(\mathbb{R}^+, \mathsf{X}), \eta([s,t]) \subset \mathsf{O}\}$,

$$\sup_{s \leq u \leq t} \left| \eta(u) - \eta(s) - \int_s^u h \circ \eta(v) \, dv \right| = 0, \quad \mathbb{Q}_x^\beta - p.s.$$

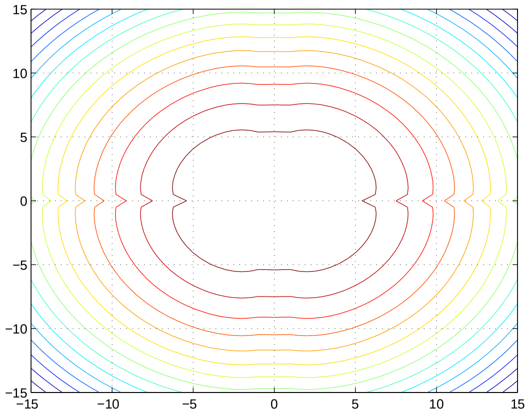
avec $$h(x) = |x|^{-\beta} \Delta_\infty(x).$$

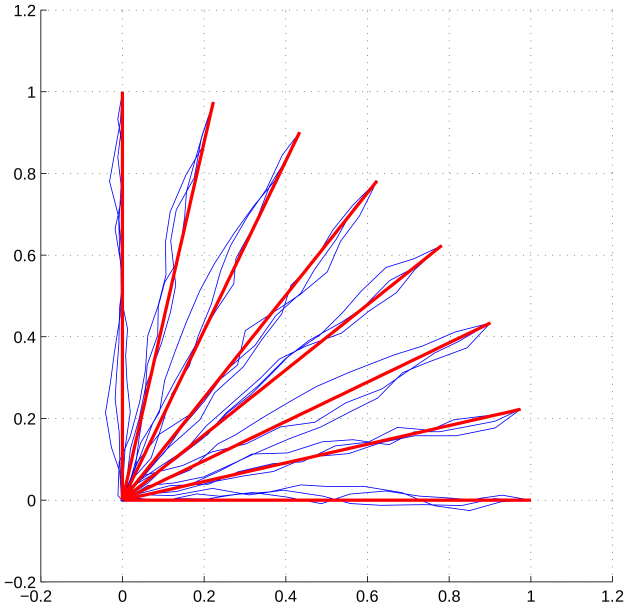
Càd Si $r^\beta \Delta(rx) \to |x|^{-\beta} \Delta_\infty(x) =: h(x)$ uniformément sur les compacts d'un ouvert O, alors la limite fluide $\mathbb{Q}_x^\beta$ charge les fonctions continues qui vérifient

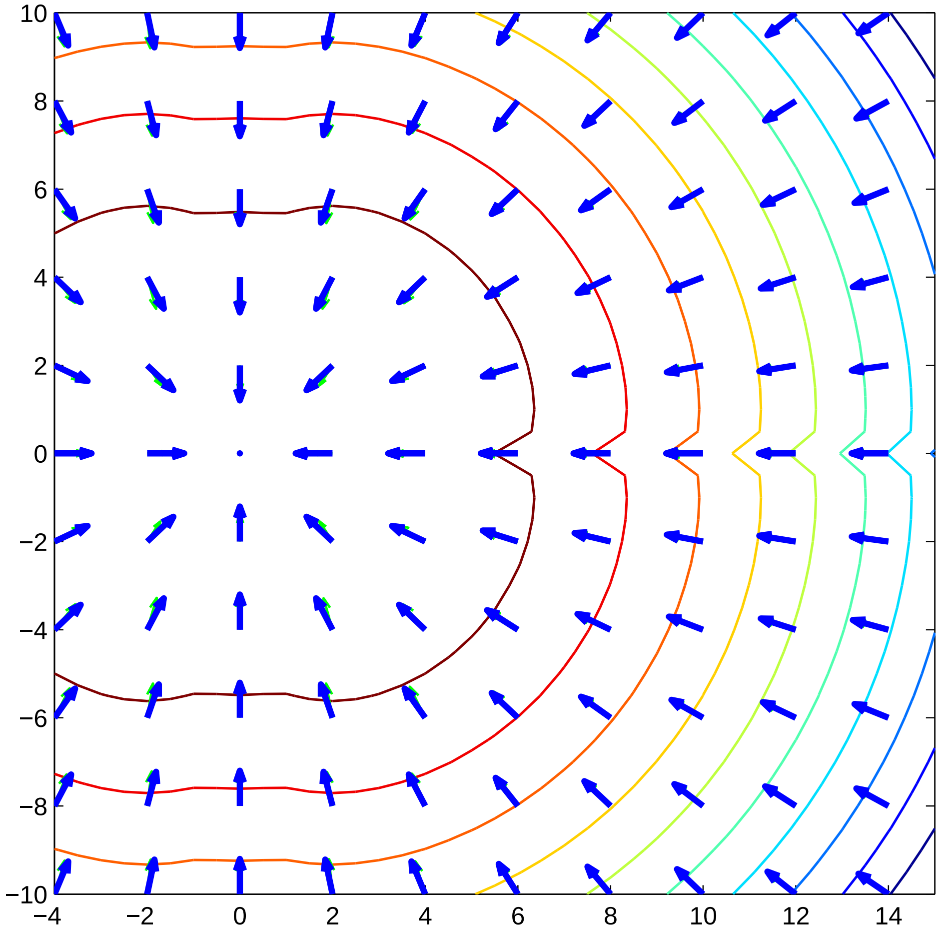
$$\eta(u) = \eta(s) + \int_s^u h(\eta(v)) dv, \qquad s \leq u \leq t,$$

dès lors que $\eta([s,t]) \subset \mathsf{O}$.

Exemple 1 $\beta = 0$ et O = X \ $\{0\}$







DHaut - Lignes de niveau, Δ et Δ_∞

GHaut - Lignes de niveau de π

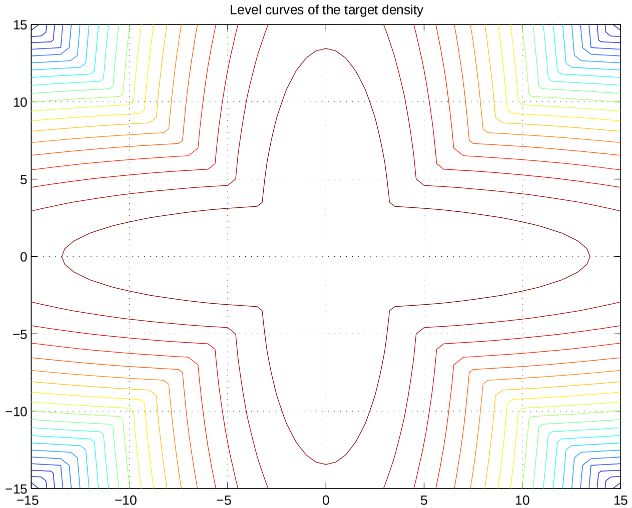
GBas - Processus η_x^β et flot de l'ODE $\dot\mu = h(\mu)$

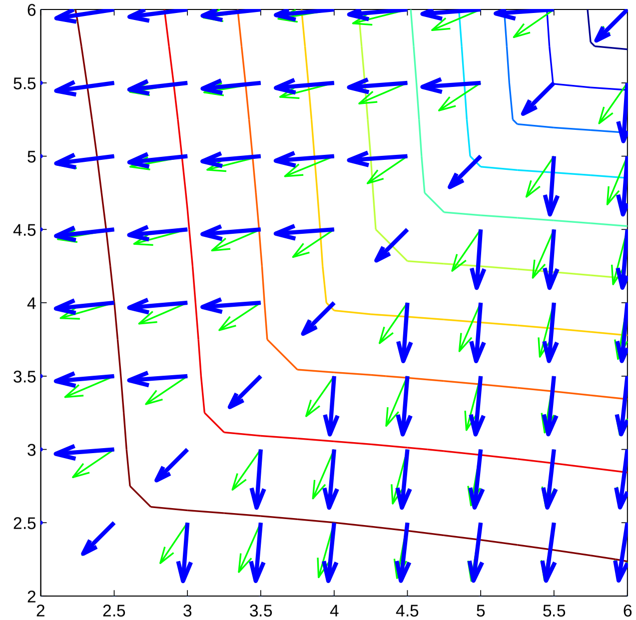
ici $h(x) = -\sigma/\sqrt{2\pi}\ x/|x|$

donc $\mu(t) = |x| - \sigma/\sqrt{2\pi}\, t$ si $\sigma t \leq \sqrt{2\pi}|x|$

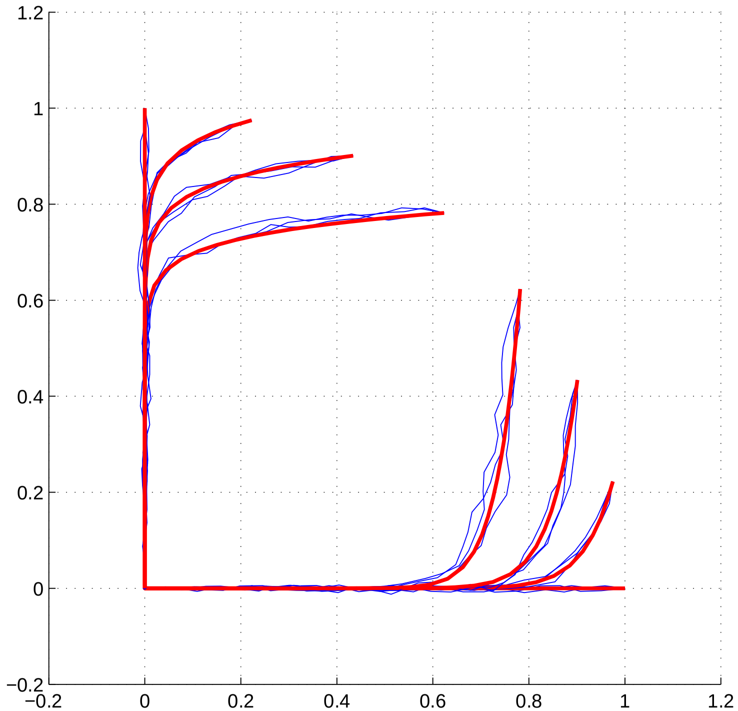
Exemple 2 π : Mélange de gaussiennes

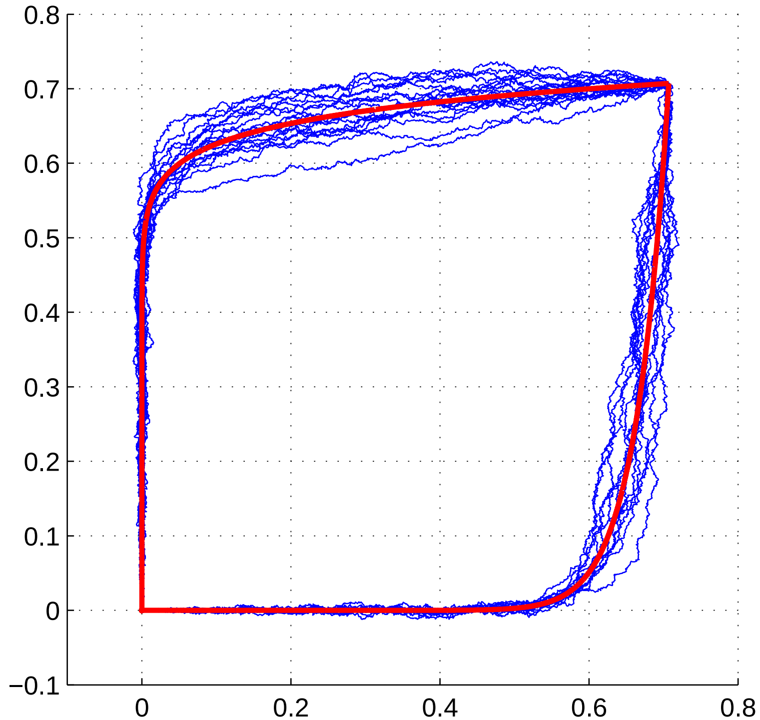
$\beta = 0$ et O $=$ X $\setminus \{|x_1| = |x_2|\}$





Δ et Δ_∞





η_r^β et Flot ODE

- Stabilité de la chaîne de Markov

Pour un taux ρ sous-géom, $f \geq 1$, $\mathbb{E}_x\left[\sum_{k=0}^{\tau_C-1} \rho(k) f(\Phi_k)\right] \leq V(x)$

$$\Uparrow$$

Sur C^c ("compact") :

$\mathbb{E}_x[V(\Phi_1)] \leq V(x) - \phi \circ V(x)$

$$\mathbb{E}_x[V(\Phi_\tau)] \leq V(x) - \mathbb{E}_x\left[\sum_{k=0}^{\tau-1} \phi \circ V(\Phi_k)\right]$$

$$\Uparrow$$

$$\mathbb{E}_x[V(\Phi_\tau)] \leq V(x) - (1-\lambda)V(x)$$
$$\mathbb{E}_x\left[\sum_{k=0}^{\tau-1} \phi \circ V(\Phi_k)\right] \leq (1-\lambda)V(x)$$

$$\Uparrow$$

τ temps d'arrêt : τ_{Φ_0} *state dependent drift criteria*

- $\tau_{\Phi_0} \longleftrightarrow \inf\{k, |\Phi_k| \leq \lambda|\Phi_0|\}, \qquad 0 < \lambda < 1$

  $\longleftrightarrow |\eta_r^\beta(\cdot, x)| \leq \lambda |x|$
- $V(x) = |x|^p$

- Stabilité de la limite fluide $\exists\ T > 0, 0 < \lambda < 1, \forall x \in \mathsf{S}_1$ sphère unité,

$$\mathbb{Q}_x^\beta\left(\eta, \inf_{[0,T]} |\eta(t)| \leq \lambda\right) = 1$$

Théorème : Conditions pour la stabilité des limites fluides?

Si

- Il existe T_0 tq pour tout $x \in \mathsf{S}_1$, toute limite fluide $\mathbb{Q}_x^\beta$

$$\mathbb{Q}_x^\beta\left(\eta, \eta([0,T_0]) \cap \mathsf{O} \neq \emptyset\right) = 1.$$

- $\forall K > 0$, $\exists T_K, 0 < \lambda_K < 1$, pour tout $x \in \mathsf{O}$, $|x| \leq K$

$$\inf_{[0,T_K \wedge T_x]} |\mu(\cdot; x)| \leq \lambda_K.$$

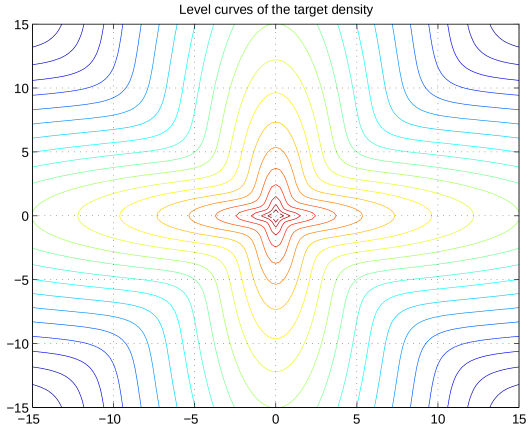
- pour tout compact $\mathsf{H} \subset \mathsf{O}$ et tout K, $\{\mu([0, T_x \wedge T_K], x), x \in \mathsf{H}\}$ est un compact de O.

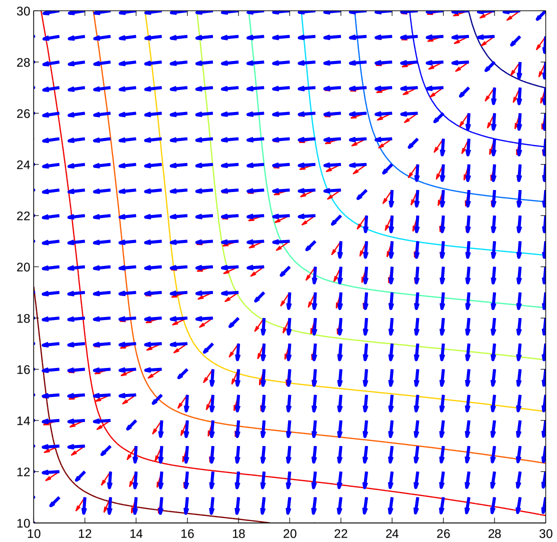
Alors

Les limites fluides sont stables.

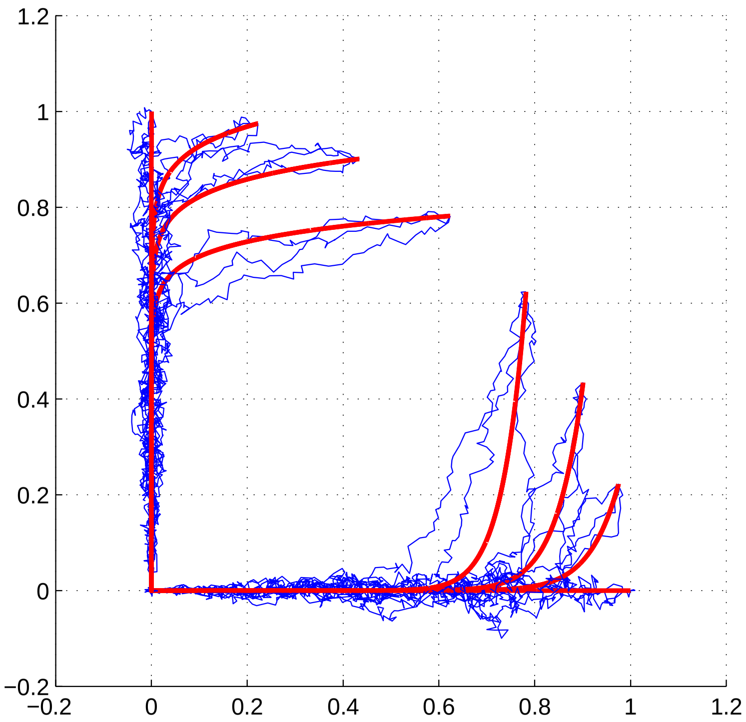
Exemple 3 π : Mélange de lois exponentielles

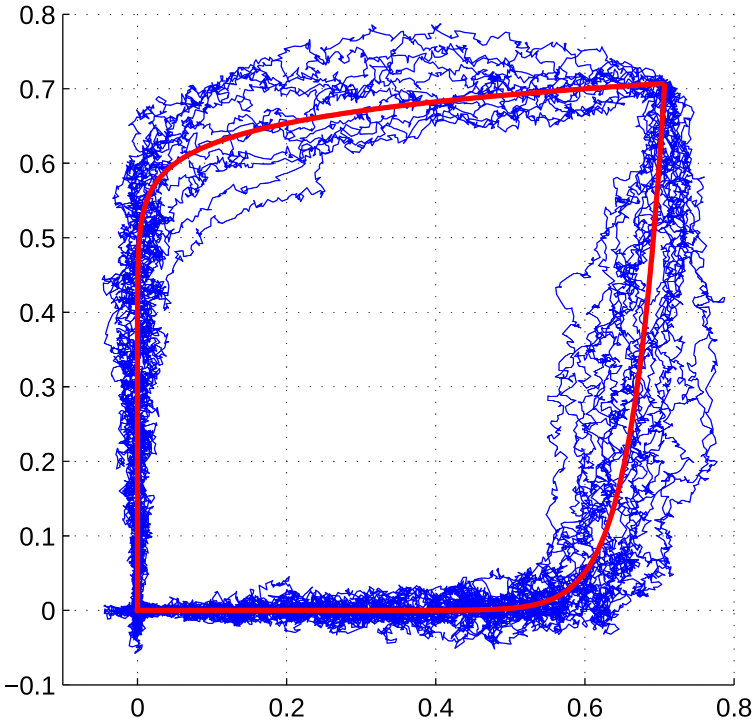
$\beta = 1 - 2\delta$ et O $=$ X $\setminus \{|x_1| = |x_2|\}$





Δ et Δ_∞





η_r^β et Flot ODE

Changement d'échelle en temps : $\eta_r^{\beta}(t,x) = \frac{1}{r}\Phi_{\lfloor tr^{1+\beta}\rfloor}$.

- Si en temps, $r^{1+\alpha}$ avec $0 \leq \alpha < \beta$

$$\mathbb{Q}_x^{\alpha} = \delta_x$$ rien ne bouge, observations trop rapprochées

- Si “bonne” vitesse : $\sup_x (1+|x|^{\beta})\Delta(x) < \infty$ où $\Phi_{k+1} = \Phi_k + \Delta(\Phi_k) + \epsilon_{k+1}$

plus β grand, plus lent est le rappel vers l'ensemble C

- Sur la limite fluide :

plus β grand, plus le temps d'entrée dans une boule de rayon $\lambda < 1$ est long.

Théorème : Sous conditions, $\lim_n \; n^{\frac{p}{1+\beta}-1} \|P^n(x,\cdot) - \pi(\cdot)\|_{\mathrm{TV}} = 0.$

Limites fluides : alternative aux méthodes de drift (quand celles-ci ne peuvent s'appliquer).

La méthode s'applique : aux chaînes *skip-free*

$$\sup_{x \in \mathsf{X}} \mathbb{E}_x[\Phi_1 - \Phi_0] < \infty.$$

`Preprint sur arXiv` :

ODE methods for skip-free Markov Chains stability with applications to MCMC. Rapport de Recherche arXiv math.PR/0607800, 2006.